

EVALUATION OF TWO COGNITIVE ABILITIES TESTS
IN A DUAL-TASK ENVIRONMENT

Michael A. Vidulich
NASA-Ames Research Center
Mail Stop 239-3
Moffett Field, CA 94035

and

Pamela S. Tsang*
NASA-Ames Research Center
Mail Stop 239-21
Moffett Field, CA 94035

ABSTRACT

Most real-world operators are required to perform multiple tasks simultaneously. In some cases, such as flying a high-performance aircraft or trouble-shooting a failing nuclear power plant, the operator's ability to "time-share" or "process in parallel" can be driven to extremes. This has created interest in selection tests of cognitive abilities. Two tests that have been suggested are the Dichotic Listening Task and the Cognitive Failures Questionnaire. Correlations between these test results and time-sharing performance were obtained and the validity of these tests were examined. The primary task was a tracking task with dynamically varying bandwidth. This was performed either alone or concurrently with either another tracking task or a spatial transformation task. The results were: (1) An unexpected negative correlation was detected between the two tests. (2) The lack of correlation between either test and task performance made the predictive utility of the tests scores appear questionable. (3) Pilots made more errors on the Dichotic Listening Task than college students.

INTRODUCTION

Many complex operational tasks, such as flying high-performance aircraft, air-traffic control, or controlling a nuclear power plant in an emergency, can be very unforgiving of errors. Therefore, it is highly desirable that the operators in charge of such tasks be as unlikely to commit an error as possible. Traditionally, the operator's training was expected to minimize error probability. However, training

* National Research Council Research Associate at NASA Ames Research Center

alone is often not the most cost-effective solution. Most notable is the problem that some people seem to be less able to learn a task than others. This is evident from the high wash-out rate found in many training programs. The resources spent on individuals who ultimately do not finish the training program are unavailable to those that do. Consequently, it is highly desirable to identify individuals who are likely to successfully complete the training program before training commences.

Of course, there are many factors that could be involved in failing to complete a training program. A few obvious examples are poor motivation, inadequate sensory acuity, inability to cope with stress, or insufficient cognitive capacity. Motivation is difficult to test in a laboratory, but since many of the jobs that have high wash-out rates are highly sought after, it seems likely that the typical trainee is well motivated. Sensory acuity can generally be measured quite accurately to insure that trainees meet an acceptable level. In general then, the most pressing need appears to be in the identification of individual differences in cognitive capacities and ability to cope with stress.

In view of the fact that cognitive control is likely to be related to performance on complex tasks, the present paper examines the relationship between time-sharing performance and two tests of cognitive abilities that have been proposed: the Dichotic Listening Task and the Cognitive Failures Questionnaire. The Dichotic Listening Task was developed by Gopher and Kahneman (1971) and is intended to test how well individuals can focus and switch attention to dichotic stimuli (i.e., different auditory stimuli simultaneously presented to each ear). The Dichotic Listening Task score (error) has been found to correlate negatively with success in flight training, to discriminate between transport pilots and fighter pilots (Gopher, 1982), and to correlate with accident proneness in bus drivers (Kahneman, Ben-Ishai, & Lotan, 1973). The Cognitive Failures Questionnaire (Broadbent, Cooper, FitzGerald, & Parkes, 1982) is a series of questions concerning the frequency of failures in perception, memory, and motor function. Through their own research as well as reviews of other's, Broadbent et al. found that the Cognitive Failures Questionnaire score is fairly stable over time. More relevant to the present paper is their finding that the various kinds of failures (i.e., perceptual, memory, or motor) all seem to occur in the same person and need not be treated as separate categories. Broadbent et al. argued that this would support the notion of some deficiency existing in overall cognitive control and that the Cognitive Failures Questionnaire score seemed to be a measure of a general likelihood of failures. So far, they have not found any significant relationships between the Cognitive Failures Questionnaire score and short-term memory, long-term memory, or dual-task performance. What they did find, suggested that the Cognitive Failures Questionnaire score would be a good indicator of how resistant an individual is to stress.

In the present experiment the two tests were administered to a group of pilots and a group of students. Performance measures on a variety of single- and dual-tasks at various levels of difficulty were

obtained. Discussion of the results will focus on the degree to which scores on the two tests are related to each other and the extent to which either test's scores is related to the single-task and dual-task performance. Inasmuch as both tests had been found to be related to the tendency of an individual to commit errors, it was expected that scores on the Cognitive Failures Questionnaire and the Dichotic Listening Task would be positively correlated. This expectation was based on the assumption that the attentional abilities evaluated by the Dichotic Listening Task might underly the "overall cognitive control" postulated by Broadbent et al. as responsible for the Cognitive Failures Questionnaire scores.

Both tests were also expected to correlate with performance, especially time-sharing performance. Correlations between good cognitive abilities, as tested by these procedures, and single-task performance would not be a problem in and of itself. But since the dual-task trials employed in the present experiment involved dynamically changing difficulty in a high workload task (and hence a potentially stressful situation), it was expected that the propensity towards cognitive failure or attentional misdirection evaluated by the tests would be manifested in the dual-task performance scores. Individuals with better (lower) scores on these tests were therefore expected to show better time-sharing performance on the experimental tasks. Although, Broadbent et al. (1982) did not obtain any correlations between the Cognitive Failures Questionnaire score and dual-task performance, a replication seemed justified. First, very little procedural detail was provided in Broadbent et al.'s review. Second, the results from the present experiment could be compared with another objective measure - that provided by the Dichotic Listening Task.

METHOD

Subjects

Twenty-four male subjects served as paid participants. Half of the subjects were pilots (with an average age of 28.8 years) and half were college students (with an average age of 21.3 years). All but two of the pilots were instrument rated and all but one had a commercial pilot's license, an instructor pilot's license, or both. Total flight time for the pilots varied from 120 hr to 2000 hr with a mean of 863 hr.

Apparatus

The experimental tasks were implemented on a PDP 11/34 minicomputer. Visual displays were presented on a CRT screen in front of the subjects and auditory stimuli were presented through stereo headphones. A joystick was mounted on the right armrest of the chair. Either another joystick or a set of eight microswitches arranged in a circle could be mounted on the left armrest. Subjects' vocal responses were processed via a Votan speech recognition device.

Tasks

Two basic tasks were used in this experiment: a tracking task and a transformation task. The tracking task was a one-dimensional compensatory tracking task with first-order control dynamics. Three levels of constant bandwidth were used: .3 Hz, .5 Hz, and .7 Hz. Also, the bandwidth could vary dynamically within a trial; ranging from .3 Hz to .7 Hz. In a given trial, the right-hand tracking task could be any one of the four levels (i.e., .3 Hz, .5 Hz, .7 Hz, or variable within the trial), but the left hand tracking always had a constant .5 Hz bandwidth.

The second task was a spatial transformation task. Stimuli designating one of eight compass directions (north, northeast, east, etc.) were presented one at a time. Subjects were required to respond with the next direction in a clockwise direction. The initial direction could be indicated either visually by the appearance of a tick mark on the CRT or auditorily by a tone of specific pitch and channel (ear). The subjects' responses could be either manual, via the microswitches on the left armrest, or vocal, via the voice recognition device. Therefore, the transformation task could be presented in any one of four possible input/output (I/O) configurations: visual/manual (VM), auditory/manual (AM), visual/speech (VS), or auditory/speech (AS).

The tracking task and the transformation task were first performed as single-tasks. In the dual-task conditions, the right-hand tracking tasks (either .5 Hz or variable bandwidth) was paired with either the left-hand tracking or one of the four transformation tasks. After some initial single- and dual-task training, a secondary task technique was adopted and the right-hand task was designated as the primary task. Subjects were instructed to maintain the primary task performance constant at the single-task level. This was to be achieved by allocating the appropriate amount of resources to the concurrent tasks according to the changes in the difficulty in the primary task. There were 10 experimental sessions. A more detailed experimental design is described in Tsang (1985).

Cognitive Ability Tests

The Dichotic Listening Task consisted of a series of 48 trials recorded on a cassette tape. Each trial consisted of two simultaneous messages, one presented to each ear. The messages were made up of simple words with a few digits embedded in each message. Each trial was divided into two sections. The first section of the trial was intended to evaluate the ability to focus attention; the second section the ability to switch attention. In section I, the subject's task was to focus on the ear indicated by a tone. Upon detecting any digits in the appropriate ear, the subject wrote it down on the appropriate line of a prepared form. A second tone indicated the beginning of section II. The subject was required to switch attention to the other ear if the second tone was different from the first. The subject's task was to record the

digits presented to the relevant ear throughout the two sections of the trial. Any deviations from the correct sequence were recorded and categorized as omissions and intrusions in the first section and as switching errors in the second section. The two sections of each trial were scored separately. Poor performance was indicated by a high total error score. The first 12 trials of the 48 run were not included in the final scores, because several subjects showed extreme practice effects during this period. The stimulus tape was an English version (Braune & Wickens, 1983) of the original Hebrew Dichotic Listening Task (Gopher & Kahneman, 1971). The entire tape, including the instructions, took approximately 35 min to complete.

The Cognitive Failures Questionnaire was taken directly from Broadbent et al. (1982). The questionnaire consisted of 25 questions describing common cognitive failures most people experience (e.g., "Do you find you confuse right and left when giving directions?"). On a five-point scale of frequency, ranging from Very Often (4) to Never (0), the subject simply circled a response to indicate how often the described event had happened in the previous 6 months. The frequency score for each response was totaled to generate the subject's score. A high frequency of cognitive failures was indicated by a high score. The Cognitive Failures Questionnaire was administered immediately after the subject had completed the Dichotic Listening Task. Both tests were administered individually.

RESULTS

The results of the present study will focus on three issues: (1) the correlation between the two tests, (2) the relationship between the test scores and performance on the experimental tasks, and (3) the effect of the background of the subjects on the test scores. Each of these topics will be dealt with in turn.

Inter-Test Correlations

Three Pearson's product-moment correlations were obtained for the ability tests; the inter-test correlations listed in Table 1 are the correlation between the two sections of the Dichotic Listening Task (Dichotic I and Dichotic II), and each with the Cognitive Failures Questionnaire (CFQ) scores. The critical r for two-tailed test ($df = 22$, $p < .05$) is .404 (Edwards, 1984). As shown in Table 1, the only positive correlation that met this criterion was the correlation between the two sections of the Dichotic Listening Task. This implies that the ability to focus attention and the ability to switch attention may be related or that focusing attention plays an important role in both sections. The unexpected negative correlation between the Cognitive Failures Questionnaire score and the Dichotic I score was significant at .05 level and that with the Dichotic II score at .1 level. These results show that the individuals who reported themselves more likely to experience cognitive failures were able to perform the Dichotic Listening Task better.

Table 1

Inter-Test Correlations

Tests	r
Dichotic I/Dichotic II	.582
Dichotic I/CFQ	-.423
Dichotic II/CFQ	-.344

Correlations between the Test Scores and Performance

Performance measures obtained included Root Mean Square Error (RMSE) for the tracking task, reaction time (RT) and percent error for the transformation task. Decrement scores, generated by subtracting the corresponding single-task performance score from any given dual-task score, were used to the dual-task analyses. Decrement scores were used to remove the effects of difficulty differences present in the single-task conditions and to isolate the magnitude of interference caused by the performance of the concurrent task. The single-task data reported here were obtained in Session 4 (last session before any dual-tasks were introduced); dual-task data in Session 7 (last dual-task session before the secondary task technique was adopted) and Session 10 (last session of the experiment).

Correlations were performed to assess the relationship between whatever abilities that are assessed by the tests and the abilities required to perform the experimental tasks. Any use of the test scores as a predictor variable presupposes that such a relationship exists. Table 2 displays the findings of these analyses: correlations between test scores and single-task performance are on top and correlations between test scores and dual-task performance are at the bottom. A positive correlation represents better performance being associated with superior (i.e., lower) test scores. In contrast, a negative correlation indicates that more reported cognitive slips on the Cognitive Failures Questionnaire or more errors on the Dichotic Listening Task are associated with better performance. In Table 2, correlations that are significantly different from zero are marked with an asterisk (2-tailed critical r (df=22) = .404, $p < .05$).

Four significant correlations between the Cognitive Failures Questionnaire and performance were obtained. However, three of the four significant correlations were negative. The subjects who reported experiencing more frequent cognitive failures tended to perform better on the single-task trials. The sole significant positive correlation occurs with the RT decrements of the dual-task trials. None of the remaining four dual-task correlations approached significance.

The two sections of the Dichotic Listening Task both show the same trends. In neither case, is there a significant positive correlation between the test scores and any dual-task performance measures. In fact, Dichotic I correlates negatively with the transformation task's RT decrements. The only strong positive correlation is with the single-task transformation task RTs. The unexpected lack of correlation with dual-task performance is problematic. Neither test demonstrated a reliable relationship with time-sharing performance in the present experiment.

Table 2
Correlations between Test Scores and Performance

Performance Measure Type	Test Score Type		
	CFQ	Dichotic I	Dichotic II
<u>Single-Task Performance</u>			
Left-Hand RMSE	-.415*	.236	-.148
Right-Hand RMSE	-.512*	.312	-.076
Transformation RT	-.543*	.675*	.635*
Transformation % Error	-.140	.305	.028
<u>Dual Tracking Performance</u>			
Left-Hand RMSE Decrement	-.187	-.225	-.077
Right-Hand RMSE Decrement	-.008	.148	.108
<u>Transformation/Tracking Performance</u>			
Right-Hand RMSE Decrement	-.092	-.098	-.205
Transformation RT Decrement	.410*	-.420*	-.173
Transformation % Error Decrement	.156	.296	-.028

* $p < .05$.

Background Effects

Table 3 displays the mean errors on the two sections of the Dichotic Listening Task obtained from the students and the pilots separately. The students committed significantly fewer omissions or intrusions in Section I ($t(22) = 1.82, p < .05$) and made fewer errors on Section II of the Dichotic Listening Task ($t(22) = 1.64, p < 0.1$). No significant difference was found between the students and the pilots on the Cognitive Failures Questionnaire (Student Mean = 38.6; Pilot Mean = 38.2). Previous results (Tsang, 1985) also indicated that there were no substantial difference in performance between these two groups.

Table 3

Students vs. Pilots
on the Dichotic Listening Task

Section	Student Errors	Pilot Errors
I	6.08	10.92
II	1.83	4.33

DISCUSSION

There are three issues to be discussed: (1) the negative correlation between the Cognitive Failures Questionnaire and the two sections of the Dichotic Listening Task, (2) the correlations between the test scores and performance, and (3) the student vs. pilot difference in the Dichotic Listening Task score.

Negative Inter-Test Correlations

Part of the inspiration for this study arose from Broadbent et al.'s (1982) suggestion that an objective correlate of Cognitive Failures Questionnaire would be useful. Hopefully, such a correlated test would be free of the "problem of defensive unwillingness to admit error," (Broadbent, et al., 1982, p. 12). Broadbent et al. reviewed several attempts to find such a correlate. Most attempts centered around some test of memory performance; none achieved very promising results. The present study was undertaken to see if the cognitive failures reported in the questionnaire were related to a subject's attentional control capabilities as detected on the more objective Dichotic Listening Task.

Surprisingly, not only were the correlations not significantly positive, they tended to be negative. These results caution against relying heavily on either of these tests as a selection tool or a classification criterion of performance on complex tasks. Replications of these results will, of course, be required and if negative correlations persist, reinterpretation of one or both tests may be unavoidable.

Correlations between Test Scores and Performance

The general paucity of positive correlations between the test scores and the performance measures is troublesome. It is tempting to explain the overall lack of positive correlations in this experiment as a result of insufficient statistical power to detect small, but

important, effects. However, this explanation does not account for the disturbing presence of the negative correlations. Nor does it account for the fact that the strongest correlations obtained with the test scores were with various measures of single-task performance.

Both tests had been expected to correlate better with the dual-task performance measures. The Dichotic Listening Task was expected to correlate better with dual-task performance because the continuous control of attention allocation was believed to be a major determinant of the dual-task performance in the present experiment. However, it is conceivable that the mechanism required for continuous attention division may be independent from attention switching. The latter being postulated to be highly related to the Dichotic Listening Task score. The Cognitive Failures Questionnaire was expected to correlate better with dual-task performance because the dual-task conditions were expected to induce higher levels of stress. But as in Broadbent et al.'s findings, no significant relationship between the Cognitive Failures Questionnaire score and dual-task performance was obtained here.

Taken as a whole, the results of this investigation suggest that the utility of these tests as predictor variables of performance in dual-task laboratory research is quite limited. It is possible that the Dichotic Listening Task will correlate with other tasks which emphasizes the switching of attention rather than its sharing. However, the present findings suggest that the predictability of the Dichotic Listening Task scores on dual-task performance may be highly task specific.

Student/Pilot Differences

One possible explanation for the difference between the students and the pilots on the Dichotic Listening Task may concern the pilots' hearing. It is possible that the pilots' hearing may have been suboptimal due to exposure to the noisy aviation environment. Whatever the explanation, the present finding suggests the possibility that experience as a pilot may be disruptive to good performance on the test. The implication is that caution must be exercised when the Dichotic Listening Task is used as a pilot trainees selection tool, especially when the pool of applicants have different levels of piloting experience and possibly various degrees of hearing damage. Again, this is a result that requires replication and careful consideration before application of the test should be taken for granted.

REFERENCES

- Braune, R. J. & Wickens, C. D. (1983). The functional age profile: An objective decision criterion for the assessment of pilot performance capacities and capabilities. In R. S. Jensen (Ed.), Proceedings of the Second Symposium on Aviation Psychology (pp. 437-444). Columbus: Ohio State University Aviation Psychology Laboratory.

- Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. British Journal of Clinical Psychology, 22, 1-16.
- Edwards, A. L. (1984). An Introduction to Linear Regression and Correlation (2nd edition). New York: Freeman.
- Gopher, D. (1982). A selective attention test as a predictor of success in flight training. Human Factors, 24, 173-183.
- Gopher, D. & Kahneman, D. (1971). Individual differences in attention and the prediction of flight criteria. Perceptual and Motor Skills, 33, 1335-1342.
- Kahneman, D., Ben-Ishai, R., & Lotan, M. (1973). Relation of a test of attention to road accidents. Journal of Applied Psychology, 58, 113-115.
- North, R. A., & Gopher, D. (1976). Measures of attention as predictors of flight performance. Human Factors, 18, 1-14.
- Tsang, P. S. (1985). Can pilots time-share better than non-pilots. In Proceedings of the Third Symposium on Aviation Psychology. Columbus: Ohio State University Aviation Psychology Laboratory.